EXTENSION OF ENSEMBLE KALMAN FILTERING
TO FOUR DIMENSIONS

T. Sauer,* B.R. Hunt, J.A. Yorke, A.V. Zimin, E. Ott, E.J. Kostelich,† I. Szunyogh, G. Gyarmati, E. Kalnay, D.J. Patil
University of Maryland, College Park, MD 20742

Abstract

We discuss a 4-dimensional ensemble Kalman filter method (4DEnKF), which is implemented in a way that observations that are asynchronous with the analysis cycle may be assimilated naturally. Tests using the Lorenz 40 variable model compare analysis error of 4DEnKF with other, ad hoc means of handling asynchronous observations.

1. INTRODUCTION

The Kalman filter is a key feature of data assimilation in numerical weather and ocean forecasting (Kalnay, 2003). Its adaptive nature is the cornerstone of efforts to optimize state estimation on the basis of a physical model and measured observations of related variables. The Kalman filter is an optimal solution in the case of linear system dynamics.

The difficulty of data assimilation in numerical weather forecasting is tied to the large number of variables in the model and the nonlinearity of the dynamics being modelled. The method of ensemble Kalman filters (EnKF) has developed as a means of attacking both problems. An ensemble of background vector fields is integrated by the model and used to estimate the current covariance matrix, a key part of the Kalman computation. Numerical experiments have shown that ensemble Kalman filters (EnKF, e.g., Evensen 1994; Evensen and van Leewen 1996, Houtekamer and Mitchell 1998, 2001; Hamill and Snyder 2000) are efficient ways to carry out data assimilation from simple models to state-of-the-art operational numerical prediction models. The ensemble square-root Kalman filter approach (Tippett et al. 2003; Bishop et al. 2001; Anderson 2001; Whitaker and Hamill 2002; Ott et al. 2002) has attracted much recent attention.

Hunt et al. (2003) proposed a means for asynchrononous observations to be assimilated in the ensemble square root Kalman filter. The Four-Dimensional Ensemble Kalman Filter (4DEnKF) is a practical way of unifying the ensemble Kalman filter and the four-dimensional variational approach. Observations can be taken into account in a natural way, even if taken at times different from assimilation times. The observational increments are propagated at intermediate time steps using the ensemble of background forecasts. This extension of the EnKF to a 4DEnKF is an analogue to the extension of the three-dimensional variational technique (3D-Var) to the four dimensional variational technique (4D-Var). The new idea is to infer the linearized model dynamics from the ensemble instead of the tangent-linear map, as done in conventional 4D-Var schemes. Furthermore, it was shown in (Hunt et al. (2003) that in the case of linear dynamics, 4DEnKF is equivalent to instantaneous assimilation of measured data.

2. ENSEMBLE KALMAN FILTERS

First recall the standard EnKF method, which assimilates observations that are time-synchronous with the analysis. Let

$$\dot{x}_m = G_m(x_1, \ldots, x_M) \tag{1}$$

for $m = 1, \ldots, M$ be a continuous dynamical system representing the background vector field, where $x = (x_1, \ldots, x_M)$. The Ensemble Kalman Filter is designed to track the evolution of an M-dimensional Gaussian distribution centered at $\overline{x}(t)$ with covariance matrix $P(t)$.

The implementation of (Ott et al. 2003; Tippett et al. 2003), follows $k+1$ trajectories of (1) starting from initial conditions $x^{a(1)}, \ldots, x^{a(k+1)}$ over a time interval $[t_a, t_b]$. We assume the system is high-dimensional, meaning that $k+1 \le M$. The $k+1$ initial conditions should be chosen so that their sample mean and sample covariance are $\overline{x}(t_a)$ and $P(t_a)$, respectively. After following the system over the time interval, denote the trajectory points at the end of the interval by $x^{b(1)}, \ldots, x^{b(k+1)}$, and compute a new sample mean $\overline{x}^b$ and sample covariance P^b from these $k+1$ vectors. This is achieved by defining the mean vector

$$\overline{x}^b = \frac{1}{k+1} \sum_{i=1}^{k+1} x^{b(i)}$$

and

$$\delta x^{b(i)} = x^{b(i)} - \overline{x}^b,$$

and defining the matrix

$$X^b = \frac{1}{\sqrt{k}} [\delta x^{b(1)} | \cdots | \delta x^{b(k+1)}]$$

*Corresponding author address: George Mason University, Fairfax, VA 22030, USA

†Current address: Arizona State University, Tempe, AZ 85287, USA

and the $M \times M$ covariance matrix

$$P^b = X^b (X^b)^T. \tag{2}$$

The maximum possible rank of P^b is k, since the sum of the columns of X^b is zero.

In the Ensemble Kalman Filter, data assimilation is done using observations assumed to have been taken at time t_b. The observations are used to replace the dynamics-generated pair $\overline{x}^b, P^b$ at time t_b with a revised pair $\overline{x}^a, P^a$ that are used as $\overline{x}(t_a')$ and $P(t_a')$ on the next time interval $[t_a', t_b']$ where $t_a' \equiv t_b$.

In the typical case the rank of P^b is k. Then the column space S of P^b is k-dimensional, and equals the row space, since P^b is a symmetric matrix. The orthonormal eigenvectors $u^{(1)}, \ldots, u^{(k)}$ of P^b that correspond to nonzero eigenvalues span this space. Since the variation of the ensemble members occurs in the directions spanning the vector space S, we generate corrections to $\overline{x}^b$ in that space. Let $Q = [u^{(1)}| \cdots |u^{(k)}]$ be the $M \times k$ matrix whose columns form a basis of S. To represent P^b in this basis, define the $k \times k$ matrix $\hat{P}^b = Q^T P^b Q$.

The data analysis step for EnKF uses observations $(y_1, \ldots, y_l)$ measured at assimilation time t_b that we assume are linearly related to the dynamical state x by $y = Hx$, where H is the observation operator. This assumption simplifies our presentation, but the method can be extended to the case of a nonlinear observation operator. Denote by R the error covariance matrix of the observations. Define $\hat{H} = HQ$ to restrict the action of H to the subspace S. Recursive weighted least squares with current solution $\overline{x}^b$ and error covariance matrix $\hat{P}^b$ yields

$$\begin{aligned} \hat{P}^a &= \hat{P}^b (I + \hat{H}^T R^{-1} \hat{H} \hat{P}^b)^{-1} \\ \Delta\hat{x} &= \hat{P}^a \hat{H}^T R^{-1} (y - H\overline{x}^b) \\ \overline{x}^a &= \overline{x}^b + Q\Delta\hat{x} \end{aligned} \tag{3}$$

The corrected most likely solution is $\overline{x}^a$, with error covariance matrix $\hat{P}^a$.

To finish the step and prepare for a new step on the next time interval, a new ensemble of $k+1$ initial conditions $x^{a(1)}, \ldots, x^{a(k+1)}$ must be produced. They should have the analysis mean $\overline{x}^a$ and analysis covariance matrix $\hat{P}^a$. One approach (Ott et al. 2002), out of many possible choices, is to define the positive square root matrix

$$Y = \{I + (\hat{X}^b)^T (\hat{P}^b)^{-1} (\hat{P}^a - \hat{P}^b)(\hat{P}^b)^{-1} \hat{X}^b\}^{1/2}, \tag{4}$$

where $\hat{X}^b = Q^T X^b$. Define the matrix $X^a = X^b Y$ and

$$X^a = \frac{1}{\sqrt{k}} [\delta x^{a(1)} | \cdots | \delta x^{a(k+1)}].$$

Next define the vectors

$$x^{a(i)} = \delta x^{a(i)} + \overline{x}^a.$$

It can be checked that

$$\begin{aligned} \overline{x}^a &= \frac{1}{k+1} \sum_{i=1}^{k+1} x^{a(i)} \\ P^a &= X^a (X^a)^T \end{aligned} \tag{5}$$

satisfy (3).

3. ASSIMILATION OF ASYNCHRONOUS DATA

The above description assumes that the data to be assimilated was observed at the assimilation time t_b. The 4DEnKF method adapts EnKF to handle asynchronous observations, those that have occurred at non-assimilation times. The key idea is to mathematically treat the observation as a slightly modified observation of the current state at the assimilation time. The method of (Hunt et al. 2003) consists of using the dynamics contained in the ensemble members to carry this out. In this way we avoid the need to linearize the original equations of motion, as is necessary in standard implementations of 4D-Var.

Notice that Eqs. (3,4,5) result in analysis vectors $x^{a(1)}, \ldots, x^{a(k+1)}$ that lie in the space spanned by the background ensemble $x^{b(1)}, \ldots, x^{b(k+1)}$. Consider model states of the form

$$x_b = \sum_{i=1}^{k+1} w_i x^{b(i)}. \tag{6}$$

The goal of the analysis is to find the appropriate set of weights $w_1^{a(j)}, \ldots, w_{k+1}^{a(j)}$ for each analysis vector $x^{a(j)}$.

Now let $y = h(x)$ be a particular observation made at time $t_c \neq t_b$. We associate to the state x_b in (6) at time t_b a corresponding state

$$x_c = \sum_{i=1}^{k+1} w_i x^{c(i)}, \tag{7}$$

where $x^{c(i)}$ is the state of the ith ensemble solution at time t_c. We assign the observation $h(x_c)$ at time t_c to the state x_b given by (6). Eqn. (7) was utilized by (Bishop et al. 2001) and (Majumdar et al. 2002) to predict the forecast effects of changes in the analysis error. Here, we use this property to propagate the dynamical information within the analysis time window.

It remains to express the asynchronous observations $h(x_c)$ as functions of x_b, the state at the analysis time. This functional relationship is needed to apply the standard recursive least squares equation as in (3). Let

$$E_b = [x^{b(1)} | \cdots | x^{b(k+1)}]$$

and

$$E_c = [x^{c(1)} | \cdots | x^{c(k+1)}]$$

be the matrices whose columns are the ensemble members at the times t_b and t_c, respectively. Then (6) and

(7) say that $E_b w = x_b$ and $E_c w = x_c$, respectively, where $w = [w_1, \ldots, w_{k+1}]^T$. The orthogonal projection to the column span of E_b is given by the matrix $E_b(E_b^T E_b)^{-1}E_b^T$, meaning that the coefficients w in (6) can be defined by $w = (E_b^T E_b)^{-1}E_b^T x_b$. The linear combination (7) is $x_c = E_c w = E_c(E_b^T E_b)^{-1}E_b^T x_b$. Therefore the observation $h(x_c)$, expressed as a function of the background state x_b at the time of assimilation, is

$$h(E_c w) = h(E_c(E_b^T E_b)^{-1}E_b^T x_b). \qquad (8)$$

The latter expression can be substituted directly into the ensemble filter equations (3). For example, a set of observations denoted by the matrix H and time-stamped at t_c can be represented at time t_b by the matrix $HE_c(E_b^T E_b)^{-1}E_b^T$. Therefore the innovation $y - Hx_c$ learned from the observations is treated instead as $y - HE_c(E_b^T E_b)^{-1}E_b^T x_b$ in the assimilation step. This technique is equivalent to the computation of the forcing of the observational increments at the correct time in 4D-Var; however, it propagates the increments forward or backward in time without the need for the linear tangent model or its adjoint.

Multiple observations are handled in the same manner. Assume the observation matrix is $H = (h_1^T | \cdots | h_r^T)^T$, where the observation row vectors $h_1, \ldots, h_r$ correspond to times $t_{c_1}, \ldots, t_{c_r}$, respectively. Then the observation matrix H in (3) is replaced with the matrix

$$\begin{pmatrix} h_1 E_{c_1} \\ \vdots \\ h_l E_{c_r} \end{pmatrix} (E_b^T E_b)^{-1}E_b^T. \qquad (9)$$

In addition, it should be noted that the t_{c_i} can be smaller or larger than t_b, allowing for observations to be used at their correct observational time even after the nominal analysis time. In the case of linear system dynamics, the 4DEnKF technique is equivalent to assimilating data at the time it is observed.

4. COMPUTER EXPERIMENTS

The differential equations model of (Lorenz, 1998) is a reasonably complex, but simply implemented spatio-temporal dynamical model that is useful for illustrating the use of 4DEnKF. Consider the vector field defined by

$$\dot{x}_m = (x_{m+1} - x_{m-2})x_{m-1} - x_m + F \qquad (10)$$

for $m = 1, \ldots, M$ and with periodic boundary conditions $x_1 = x_{M+1}$. In experiments shown here, the forcing parameter was set to $F = 8$ and the system dimension $M = 40$. Under these conditions, the system dynamics exhibit high-dimensional chaos.

A long background trajectory x^*, to be considered as the true trajectory, was integrated. The average root mean square deviation from the mean is approximately 3.61 for the true trajectory. We produced artificial noisy observations at each time interval Δt by adding uncorrelated Gaussian noise with variance 1 to the true state at each spatial location.

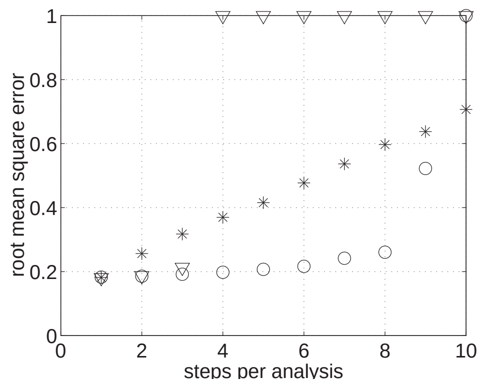

FIG. 1: Root mean square error of proposed 4DEnKF method (circles) compared to standard EnKF (asterisks) and EnKF with time interpolation (triangles). Variance inflation is set at 0.005 per time step. Symbols showing RMSE = 1 actually represent values ≥ 1. RMSE is averaged over several runs of 40,000 steps.

Figure 1 shows that if we use 4DEnKF, assimilations can be skipped with little loss of accuracy in tracking the system state. The system is advanced in steps of size $\Delta t = .05$, but instead of assimilating the observations at each step, assimilation of past data is done only every every s steps. The resulting root mean square error (RMSE) is plotted as circles in Figure 1 as a function of s. For $s \leq 6$, it appears that little accuracy is lost. The fact that the RMSE in Fig. 1 stays constant as s increases shows the ability of 4DEnKF to take asynchronous observations into account without carrying out analysis at each observation step, for the Lorenz40 example. As mentioned above, it can be shown analytically that the 4DEnKF method is equivalent to assimilating at each observation time in the case of linear background dynamics. The experiment shows that the property can hold as well for chaotic, nonlinear dynamics, at least for small values of s.

The RMSE of two other methods are shown in Fig. 1 for comparison. The asterisks denote the RMSE found by using standard EnSQKF, allowing s steps of length Δt to elapse between assimilations. Only those observations occurring at the assimilation time were used for assimilation. The triangles refer to time-interpolation of the data since the last assimilation. In this alternative, linear interpolation of individual observations as a function of the ensemble background state evolved by the model is used to create an improved observation $y_\triangle(t_b)$

at the assimilation time. In other words, the innovation at time t_c is added instead at assimilation time t_b. For the Lorenz example, where the observations are noisy states, this amounts to replacing the observation at time t_c with $y_\triangle(t_b) \equiv y(t_c) + \overline{x}_b - \overline{x}_c$ for assimilation at time t_b, which is carried out by standard EnKF. The idea behind this technique is widely used in operational 3D-Var systems to assimilate asynchronous observations (e.g., Huang et al. 2002; Benjamin et al. 2003). Our implementation provides somewhat optimistic results for this technique, since our background error covariance matrix is not static (independent of time) and homogeneous (independent of location) as it is assumed in a 3D-Var. As Figure 1 shows, for the latter two methods, the accuracy of the assimilated system state becomes considerably worse compared to 4DEnKF as the steps per assimilation s increases.

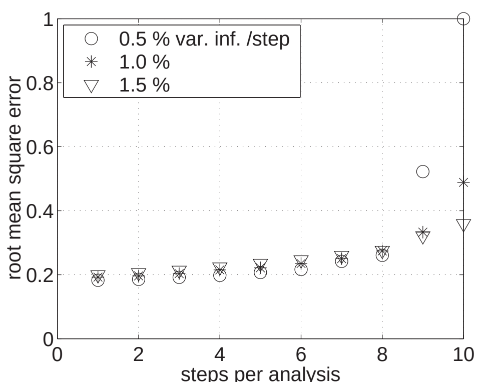

FIG. 2: Variance inflation has only a small effect on 4DEnKF. The RMSE of the method is shown for $\epsilon = 0.005$ (circles), 0.01 (asterisks), and 0.015 (triangles) per time step.

Variance inflation was used in the experiments described above, meaning that the analysis covariance matrix was artificially inflated by adding ϵI to $\hat{P}^a$ for small ϵ. In Figure 1, $\epsilon = 0.005$ per time step was used for all methods. Variance inflation helps to compensate for underestimation of the uncertainty in the background state due to nonlinearity, limited ensemble size, and model error.

Figure2 shows the effect of varying the amount of variance inflation applied to the 4DEnKF method. Variance inflation is often done by enlarging $\hat{P}^b$ rather than $\hat{P}^a$, in order to make up for model evolution that is not captured by the ensemble members. In this experiment we enlarged $\hat{P}^a$ instead, so that the ensemble would not have to be adjusted before and after the analysis step. The two approaches yielded quite similar results.

The results in Fig. 1 show that for pre-assimilation data, 4DEnKF is superior to straightforward EnKF as well as an alternative form where observational increments were computed with the background at the observing time, a method also used in operational centers. We have also achieved similar results by applying the 4DEnKF methodology to the Local Ensemble Kalman Filter (LEKF), as developed in (Ott et al. 2003). The local approach is based on the hypothesis that assimilation can be done on moderate-size spatial domains and reassembled. The 4D treatment of the asynchronous local observations can be exploited in the same way as shown in this article.

The computational savings possible with the 4DEnKF technique arise from the ability to improve the use of asynchronous observations without more frequent assimilations. The extra computational cost of 4DEnKF is dominated by inverting the $(k+1) \times (k+1)$ matrix $E_b^T E_b$ in (8), which is comparatively small if the ensemble size $k+1$ is small compared to the number of state variables M. Moreover, applying this technique in conjunction with local domains as in LEKF allows k to be greatly reduced in comparison with M.

5. ACKNOWLEDGEMENTS

This work was supported by a James S. McDonnell 21st Century Research Award, by the Office of Naval Research (Physics), by the Army Research Office, by the NPOESS Integrated Program Office, and by the National Science Foundation (Grants DMS 0104087, DMS 0208092, and PHYS 0098632).

6. REFERENCES

Anderson, J. L., 2001: An ensemble adjustment filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884-2903.

Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **127**, 2741-2758.

Benjamin, S. G. et al., 2003: An hourly assimilation/forecast cycle: The RUC. *Mon. Wea. Rev.* **132**,495-518.

Bishop, C. H., B. J. Etherton, and S. Majumdar, 2001: Adaptive sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical aspects. *Mon. Wea. Rev.* **129**, 420-436.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**(C5), 10 143-10 162.

Evensen, G., and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman Filter with a quasi-geostrophic model. *Mon. Wea. Rev.* **124**, 85-96.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman Filter-3D variational analysis scheme. *Mon. Wea. Rev.* **128**, 2905-2919.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman Filter technique. *Mon. Wea. Rev.* **126**, 796-811.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123-137.

Huang, X.-Y., K. S. Morgensen, and X. Yang, 2002: First-guess at the appropriate time: the HIRLAM implementation and experiments. Proceedings, HIRLAM Workshop on Variational Data Assimilation and Remote Sensing, Helsinki, Finland, 28-43.

Hunt, B.R., E. Kalnay, E.J. Kostelich, E. Ott, D.J. Patil, T. Sauer, I. Szunyogh, J.A. Yorke, A.V. Zimin, 2003: Four-Dimensional Ensemble Kalman Filtering. To appear, Tellus A.

Kalnay, E., 2003: *Atmospheric modeling, data assimilation, and predictability.* Cambridge University Press, Cambridge, 341 pp.

Lorenz, E.N. 1998: In Proc. Seminar on Predictability, Vol. 1 (ECMWF, Reading, Berkshire, UK 1996); E. N. Lorenz and K.A. Emanuel, 1998: *J. Atmos. Sci.* **55**, 399.

Majumdar, S., C.H. Bishop, J. Etherton, and Z. Toth, 2002: Adaptive sampling with the ensemble transform Kalman filter. Part II: Field program implementation. *Mon. Wea. Rev.* **130**, 1144-1165.

Ott, E., B. R. Hunt, I. Szunyogh, A.V. Zimin, E.J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil, J.A. Yorke, 2002: A local ensemble Kalman filter for atmospheric data assimilation. To appear, Tellus A.

Ott, E., B. R. Hunt, I. Szunyogh, A.V. Zimin, E.J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil, J.A. Yorke, 2003: Estimating the state of large spatio-temporal chaotic systems. To appear, Phys. Rev.

Tippett, M.K., J.L. Anderson, C.H. Bishop, T.M. Hamill, J.S. Whitaker, 2003: Ensemble square-root filters. Mon. Wea. Rev. **131**, 1485-90.

Whitaker, J. S., and T. H. Hamill, 2002: Ensemble Data Assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1913-1924.